Landscape Ecology **19**: 273–289, 2004.
© 2004 *Kluwer Academic Publishers. Printed in the Netherlands.*

Research article

Modelling land cover transitions: A solution to the problem of spatial dependence in data

Kevin Weaver and Ajith H. Perera*
*Forest Landscape Ecology Program, Ontario Forest Research Institute, Ontario Ministry of Natural Resources, 1235 Queen St. East, Sault Ste. Marie, Ontario, Canada P6A 2E5; *Author for correspondence (tel.: 705-946-2981 ext. 217; fax: 705-946-2030; e-mail: ajith.perera@mnr.gov.on.ca)*

Received 26 September 2002; accepted in revised form 12 May 2003

Key words: Fragmentation, Regions, Stochastic, Spatial bias, Vegetation transition

Abstract

Raster-based spatial land cover transition models (LCTMs) are widely used in landscape ecology. However, many LCTMs do not account for spatial dependence of the input data, which may artificially fragment the output spatial configuration. We demonstrate the consequences of ignoring spatial dependence, thus assigning probabilities randomly in space, using a simple LCTM. We ran the model from four different initial conditions with distinct spatial configurations and results indicated that, after 20 simulation steps, all of them converged towards the spatial configuration of the random data set. From an ecological perspective this is a serious problem because ecological data often exhibit distinct spatial configuration related to ecological processes. As a solution, we propose an approach (region approach) that accounts for spatial dependence of LCTM input data. Underlying spatial dependence was used to apply spatial bias to probability assignment within the model. As a case study we applied a region approach to a Vegetation Transition Model (VTM); a semi-Markovian model that simulates forest succession. The VTM was applied to approximately 500,000 ha of boreal forest in Ontario, at 1 ha pixel resolution. When the stochastic transition algorithms were applied without accounting for spatial dependence, spatial configuration of the output data became progressively more fragmented. When the VTM was applied using the region approach to account for spatial dependence output fragmentation was reduced. Accounting for spatial dependence in transition models will create more reliable output for analyzing spatial patterns and relating those patterns to ecological processes.

Introduction

Background

Spatial transition models based on stochastic processes (Moore 1990) in space and time at various scales (Sklar and Costanza 1991) are widely used in ecology. Raster-based models, where each pixel is at an explicit spatial location (Baker 1989), are increasing in popularity because of advances in geographic information systems (GIS). An advantage of raster-based models is that many spatial scales can be modelled at a single spatial resolution and extent (Stocks and Wise 2000). Raster-based land cover transition models (LCTMs) provide ecologists a chance to better understand landscape changes, the relationship between pattern and process, and the opportunity to simulate data that is otherwise unattainable (e.g., future landscapes). Transition models are useful when the process of change cannot be modelled as a mechanistic process (Turner 1989). Applications of LCTMs in ecology include modelling land cover change (e.g., Turner 1987; Muller and Middleton 1994; Brown et al. 2000; Jenerette and Wu 2001), grassland dynamics (e.g., Matsinos and Troumbis 2002), rangeland degradation (e.g., Li and Reynolds

1997), and forest landscape dynamics (e.g., Acevedo et al. 1996; Balzter et al. 1998; He and Mladenoff 1999; Yemshanov and Perera 2002).

Popular as these models are in predicting land cover change in space, many LCTMs do not account for spatial dependence; i.e., they simulate the fate of each pixel independently. In fact, the assumption that the future state of a pixel does not depend on the state of its neighbours has dominated ecological models (Hunsaker et al. 2001), with the exception of cellular automata (CA) models (Wolfram 1984). Models that incorporate CA theory treat time and space as discrete entities with local interactions (Sirakoulis et al. 2000). CA models (e.g., Turner 1987; Wissel 1992; Li and Reynolds 1997; Molofsky et al. 2001) apply neighbourhood influence to account for the fact that simulated data are spatially dependent.

Spatial dependence is the generalization that values closer together are more similar than values further apart (Cliff and Ord 1973; Rossi et al. 1992). The definition of spatial dependence can be extended further to include directionality. The direction of spatial dependence explains whether objects are similar or dissimilar in space. For example, similar values close together in space have positive spatial dependence; dissimilar values close together have negative spatial dependence; and, similar values independent of location have no spatial dependence (Ding and Fotheringham 1992). Here spatial dependence refers to positive spatial dependence unless otherwise stated. Spatial dependence in forest landscape ecology would include similar site types located near each other, or similar tree species occupying the same area. For example, Fox et al. (2001) provides a review of studies that observed spatial dependence in species attributes for different stages of vegetation development. Anselin (1996) suggested that spatial dependence in ecological data is the rule, rather than the exception. If that is true, it makes sense to consider spatial dependence when simulating data using LCTMs.

Spatial dependence in LCTMs

Not accounting for spatial dependence in a model may affect spatial output. If the spatial dependence of the simulated data are ignored in the simulation process output may become fragmented (e.g., Turner 1987) due to the stochasticity of the model algorithm. For example, all pixels representing a contiguous forest stand containing one cover type with homogeneous soil conditions would be expected to remain relatively aggregated through time. However, ignoring spatial dependence means that each pixel is modelled independently in space and there is no way to ensure that those spatially clustered pixels will behave similarly, which may lead to fragmentation. By not accounting for spatial dependence output may be misleading, which in turn may lead to improper ecological interpretations and decision making by model users. The problem of not accounting for spatial dependence in spatial stochastic modelling has been mentioned in the past (e.g., Barringer and Robinson 1981; Bell and Hinojosa 1977; Turner 1987; Li and Reynolds 1997), but is often given little attention in LCTM development.

The problem of not accounting for spatial dependence in the modelling process is scale dependent. For example, polygon or vector data intrinsically account for spatial dependence at least within polygon boundaries. However, it may be necessary to account for spatial dependence beyond those boundaries. With raster data, pixel resolution will determine how likely the output is to become fragmented due to stochastic simulation. For models that simulate single pixels large enough to encompass forest stands, spatial dependence is less of an issue because fragmentation at that scale would not be considered a problem. Given the same modelled process, the smaller pixels get the more susceptible they become to fragmentation as a result of not accounting for spatial dependence in LCTMs.

A possible solution for LCTMs that ignore spatial dependence is to include some form of spatial bias when assigning transitions in space so that similar land cover types have a greater chance of occurring near each other, reducing the risk of artificial fragmentation. The concept of applying spatial bias in modelling is no different than spatial biases that operate in nature. Terrain attributes (e.g., steep slope), soil properties (e.g., high moisture content), and forest cover attributes (e.g., seed dispersal) create spatial bias that influences what, when, and where forest cover types are likely to locate, leading to a natural aggregation of similar types in space. For example, Turner (1989) suggests that topography and soil attributes may bias the size and arrangement of patches. Rossi et al. (1992) stated that forest cover types often differ on south and north facing slopes. In nature, both physical and biological factors bias the formation of patches at multiple scales (Deutschman et al. 1993). Wiens (1976) identified various factors that bias patch formation including dispersal, micro-topography, soil

conditions, and climate (macro and micro). Green (1989) points out that tree species in a tropical rainforest contain clumped distributions mainly because distance of seed dispersal is limited. Similarly, Frelich and Reich (1999) found that neighbourhood effects such as seed rain, shading, and changes in nutrient availability are important drivers of the spatial dynamics of some forest cover types. Fox et al. (2001) found that microsite effects were the cause of spatial dependence of species attributes in several studies. For example, in temperate and tropical forests, edaphic conditions influence the pattern of forest species (Svenning and Skov 2002). Levin (1976) used the term 'local uniqueness' in referring to the cause of spatial heterogeneity created from the bias of many environmental factors. Similar to the way spatial bias works in nature, it can be applied in LCTMs by using complementary spatial information to encourage patch formation or aggregation and deter fragmentation that may occur without it.

Study objectives

The main objectives of this paper are to: 1) demonstrate the consequences of not accounting for spatial dependence in a LCTM and 2) demonstrate the effects of accounting for spatial dependence by applying spatial bias in a LCTM of forest landscape dynamics. We tested two null hypotheses for the first objective. The first was that there is no significant difference in spatial configurations among simulated output of four initially distinctly different data sets. The second null hypothesis was that the simulated output spatial configurations of the non-random data sets are not significantly more fragmented than the initial data set. Failing to reject the first hypothesis but rejecting the second would provide strong evidence that the problem of artificial fragmentation exists when spatial dependence is ignored. In fulfillment of the second objective as a potential solution to artificial fragmentation, we tested a third null hypothesis: that there is no significant difference in spatial configuration between output simulated with and without accounting for spatial dependence.

Methods

Spatial configuration

For this study, spatial configuration was defined as the size, shape, and arrangement of patches on the landscape. We defined landscape and patch in spatial, but not necessarily ecological, terms. A landscape is made up of spatially arranged "landscape elements" (Forman and Godron 1986), referred to as pixels in raster data. A patch then, is a contiguous cluster of pixels of the same classification on the landscape.

Many tools are available to quantify the spatial configuration of a landscape (e.g., r.le [Baker and Cai 1992], FRAGSTATS [McGarigal and Marks 1995], LEAP II [Perera et al. 1997]). Because there is no one landscape pattern index (LPI) to characterize the spatial configuration of a landscape, several LPIs are often used to get a better representation (Riitters et al. 1995; Hargis et al. 1998). The advantage of using LPIs is that the relationship between indices and spatial configuration has been well documented (e.g., Riitters et al. 1995; Li and Reynolds 1994; McGarigal and Marks 1995). In this study, a suite of eight LPIs were used to quantify the spatial configuration of simulated output: mean patch size (MPS), patch size coefficient of variation (PSCV), patch density (PD), double log fractal dimension (DLFD), mean proximity index (MPI), interspersion and juxtaposition index (IJI), mean nearest-neighbour distance (MNN), and contagion index (CONTAG). MPS represents the total landscape area divided by the number of patches. PSCV measures the relative variability (%) of patch sizes around the mean. PD is the number of patches per 100 ha. DLFD indicates the shape complexity of patches on the landscape. MPI is a measure of the distance to nearest patches of the same classification and the patch size of those neighbours. IJI is the probability of a patch being adjacent to patches of other classes. MNN is the sum of the distance to neighbouring patches of the same classification, divided by the number of neighbours. CONTAG is an indicator of how clumped or dispersed patches of the same classification are on the landscape (see McGarigal and Marks 1995 for further description of all LPIs).

All LPIs were calculated, using the LPI module of LEAP II (Perera et al. 1997), at two different spatial extents (entire and local) to provide a more robust analysis, because pattern detected at one scale may differ from pattern detected at another scale (Turner

et al. 1989, O'Neill et al. 1996, Baldwin et al. 2001). For the entire extent, each LPI was calculated across the entire landscape so that the overall spatial configuration was characterized by one value. For the local extent, multiple values for each LPI were calculated across the landscape using analysis windows 50^2 pixels in size. An algorithm was written using Arc Macro Language in ArcInfo (ESRI 1998) to place 21 non-overlapping random windows (approximately 10% of the landscape), extract, and convert the data for use in LEAP II.

Study 1: consequences of ignoring spatial dependence in LCTMs

A simple LCTM, based on the model developed by Yemshanov and Perera (2002), was used to demonstrate the consequences of ignoring spatial dependence. The LCTM contained one transition probability matrix for five discrete cover types (A, B, C, D, E). The conditional probability vector of transition from any given cover type to cover types A, B, C, D, and E was 0.03, 0.4, 0.17, 0.21, and 0.19, respectively. Every cover type had some probability of self-replacement. The model was representative of many LCTMs in that there was no mechanism to account for spatial dependence of the simulated data. For simplicity and interpretation of LPIs, conditional probabilities were conFigured so that cover type composition within each data set would remain relatively constant throughout the simulation period. The transition process began by sequentially searching the data set until a pixel eligible for transition was found. The transition probability vector for the cover type of that pixel was transformed into a cumulative range from zero to one. Based on where a uniformly distributed pseudo-random number fell in the cumulative range, a transition to one of the five cover types was assigned to that pixel. The process repeated until all pixels had been assigned a transition, at which point the simulation step ended and the next simulation step began.

Four raster data sets, each 725^2 pixels, were created as input into the transition probability model. Each data set contained five cover types of unequal proportions (A = 3%, B = 40%, C = 17%, D = 21%, E = 19%) although the proportion of area occupied by any given cover type was the same for all four data sets. The data sets were created to have distinct spatial configurations ranging from completely random to large contiguous blocks. We refer to the raster data sets as random, small patch, large patch, and block (further described in Figure 4). The random data set was created using RULE (Gardner 1999) and the three non-random sets were created using various spatial filters in ArcInfo (ESRI 1998). The purpose of having data sets with a range of spatial configurations was to examine how simulating those data sets, without accounting for spatial dependence, would affect their spatial configuration.

The transition probability model was applied for 25 simulation replicates for each data set for a period of 100 simulation steps. LPIs were used to quantify the spatial configuration for each simulated output at three simulation steps (20, 60, and 100) and two extents (entire and local). For the local extent, values for the 21 windows were averaged for each simulation replicate prior to analysis. A one-way analysis of variance (ANOVA) design (SPSS 1998) was used to test for significant ($\alpha = 0.05$) differences among simulated output spatial configuration for the study data sets ($n = 100$) corresponding to the first null hypothesis. The ANOVA was conducted for LPIs as the dependent variable at two extents and three simulation steps. Duncan's post-hoc test ($\alpha = 0.05$) with harmonic means (SPSS 1998) was used to determine which combination of data sets differed. Multiple *t*-tests, with one-sided alternative hypotheses (Table 1), were used to compare spatial configuration between the mean of sample output at simulation step 20 ($n = 25$) and the initial data set ($n = 1$) in relation to the second null hypothesis. For the *t*-test, significance ($\alpha = 0.05$) was tested for three data sets (excluding the random one) for seven LPIs (excluding MPI).

Study 2: accounting for spatial dependence in LCTMs

A Vegetation Transition Model (VTM) was used to demonstrate the effects of accounting for spatial dependence. The VTM is a sub-module of the Boreal Forest Landscape Dynamics Simulator (BFOLDS) Model (Perera et al. 2003) that simulates forest cover change, using dominant forest cover types, at a temporal resolution of one year (further referred to as a simulation step) based on transition probability matrices. Transition probability matrices were developed as semi-Markov chains from published studies on boreal forest succession (Yemshanov and Perera 2002). Holding times for the semi-Markov chains were based on a Gamma probability distribution function (Yemshanov and Perera 2002). The beta version of

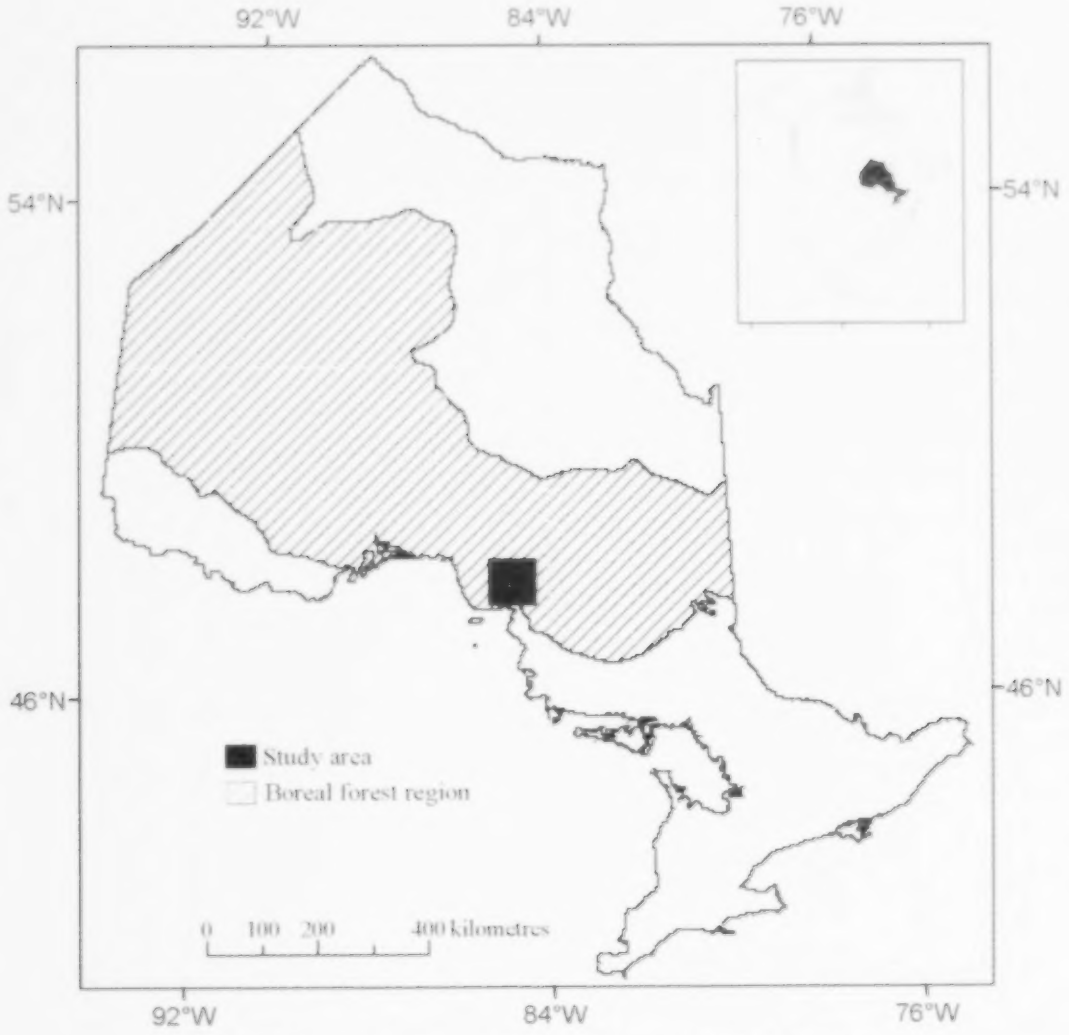

Figure 1. Location of study area in the boreal forest region of Ontario, Canada.

the VTM was parameterized to simulate succession in the boreal forest of Ontario and does not account for spatial dependence of input data.

The transition probability matrices are stratified by geoclimate, field capacity, soil nutrient status, and age of cover type and probabilities are based on the cover type of a pixel and the time spent as that cover type. Each transition probability matrix contains probabilities, between zero and one, of moving from one cover type at time t to another cover type at time after holding time has elapsed ($t+1$). A total of 15 forest cover types were defined for the boreal forest of Ontario. Assignment of transitions in space is handled one pixel at a time in the VTM using supplementary information at each pixel's location to determine the appropriate transition probability matrix. The VTM requires raster input data at 1 ha resolution and simulates all processes at that resolution. For more details on the VTM and BFOLDS see Yemshanov and Perera (2002) and Perera et al. (2003).

Model inputs include spatial layers of cover type, age of cover type, field capacity, soil nutrient status, and geoclimatic zone. Cover type and associated age layers were created from Ontario Forest Resource In-

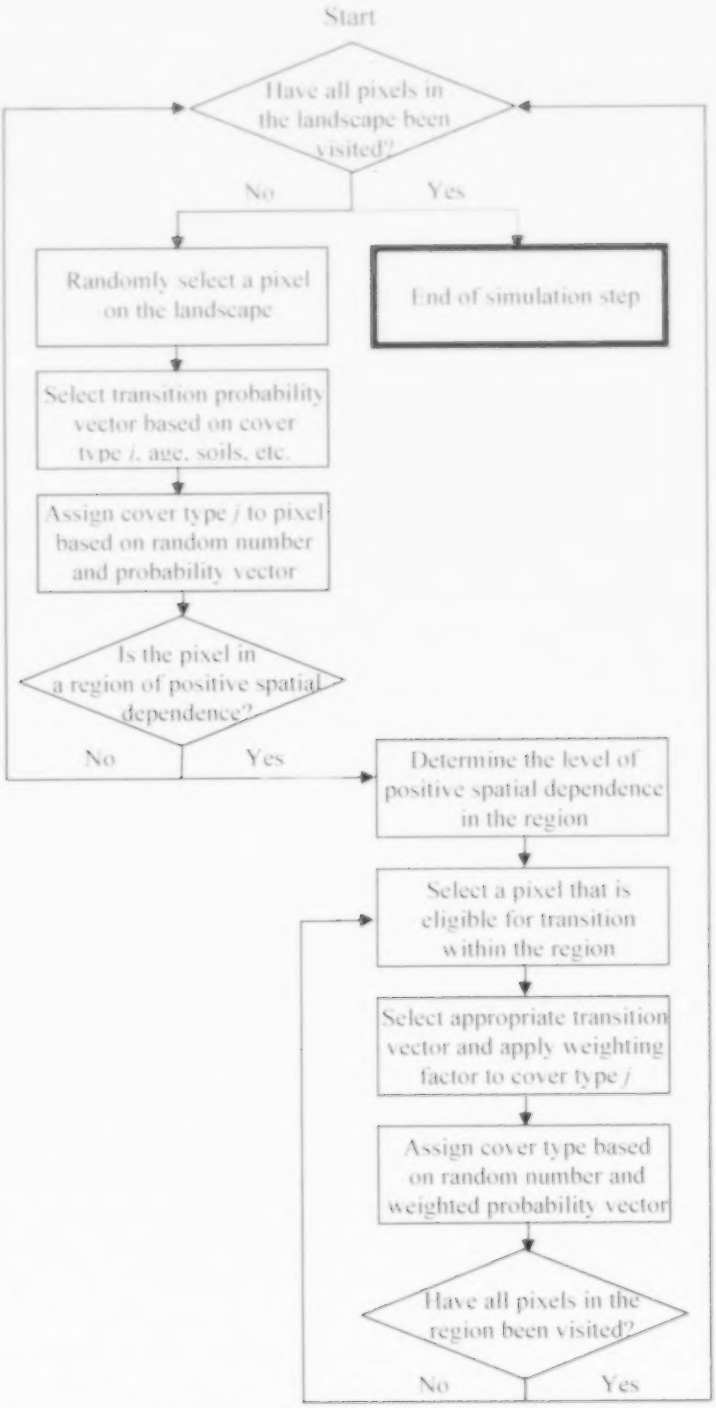
Start
Have all pixels in the landscape been visited?
No
Yes
Randomly select a pixel on the landscape
End of simulation step
Select transition probability vector based on cover type i, age, soils, etc.
Assign cover type j to pixel based on random number and probability vector
Is the pixel in a region of positive spatial dependence?
No
Yes
Determine the level of positive spatial dependence in the region
Select a pixel that is eligible for transition within the region
Select appropriate transition vector and apply weighting factor to cover type j
Assign cover type based on random number and weighted probability vector
Have all pixels in the region been visited?
No
Yes

Figure 2. Algorithm for integrating spatial bias into the assignment of transition probabilities in the Vegetation Transition Model.

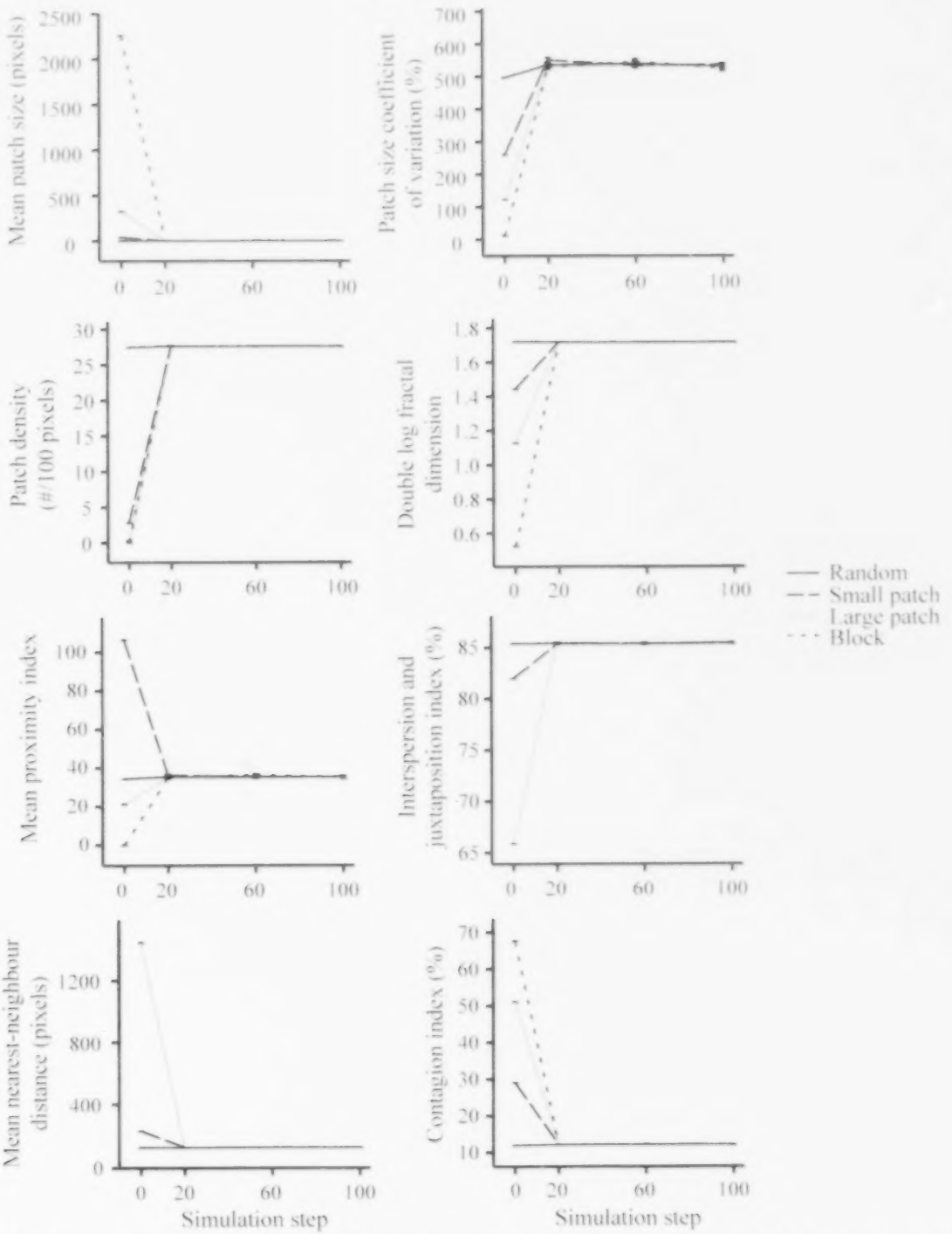

Figure 3. Means (± SE, n = 25) of landscape pattern indices (measured at the local extent) for the simulated study data sets. Note: Interspersion and juxtaposition index and mean nearest-neighbour distance could not be calculated for the block data set at simulation step 0 because only one class type occurred in the local windows.

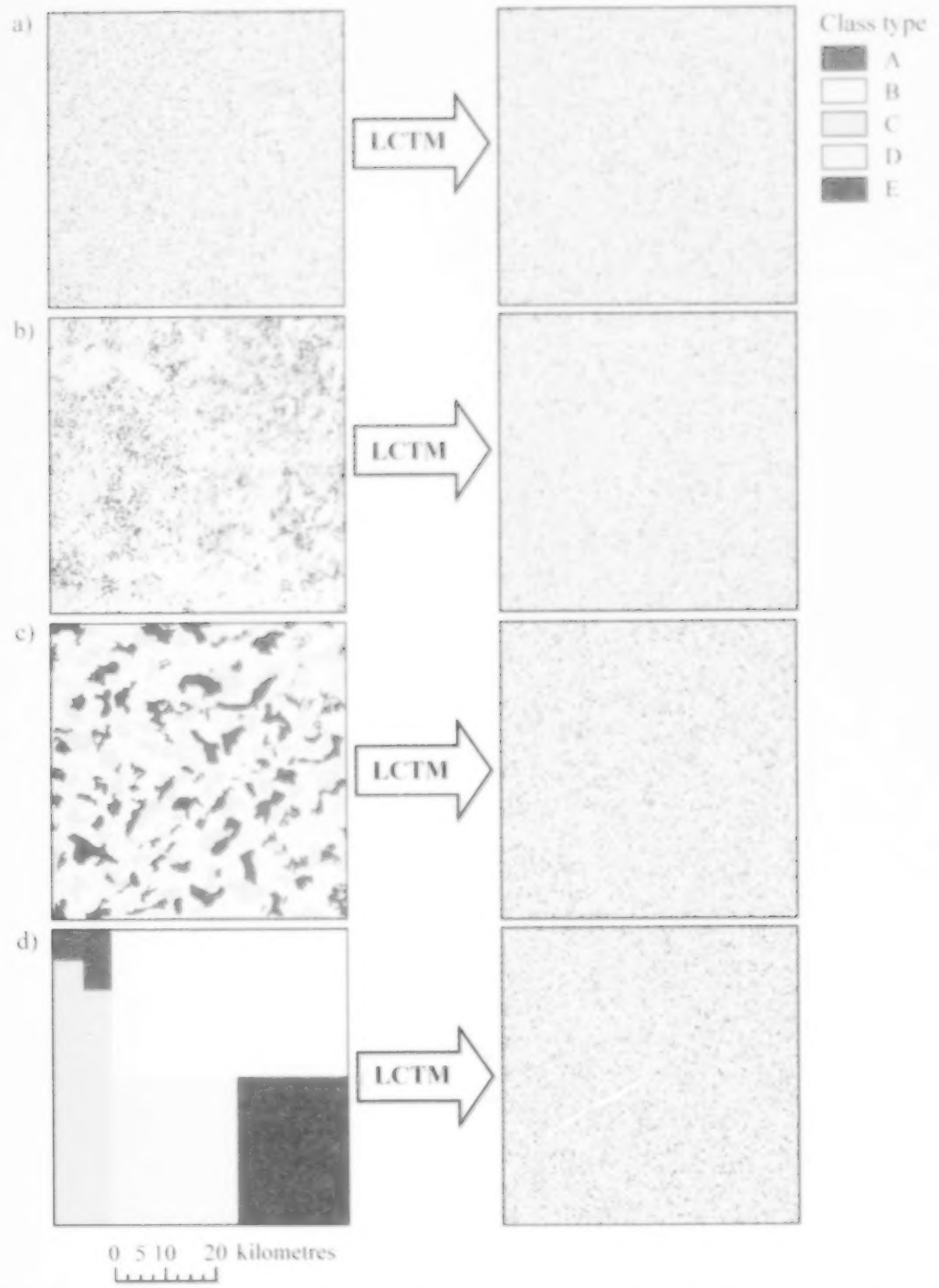

Figure 4. Initial raster data sets (725² pixels) with distinct spatial configurations: a) random, b) small patch, c) large patch, and d) block and corresponding output at simulation step 20, simulated using a land cover transition model (LCTM). Note: Class type composition (A = 3%, B = 40%, C = 17%, D = 21%, E = 19%) was the same for all initial data sets.

Table 1. Null and one-sided alternative hypotheses for t-tests comparing spatial configuration between the initial data and output at simulation step 20

Landscape pattern index	H_0	H_1
Mean patch size	$\mu_{s1} = \mu_{s0}$	$\mu_{s1} > \mu_{s0}$
Patch size coefficient of variation	$\mu_{s1} = \mu_{s0}$	$\mu_{s1} > \mu_{s0}$
Patch density	$\mu_{s1} = \mu_{s0}$	$\mu_{s1} < \mu_{s0}$
Double log fractal dimension	$\mu_{s1} = \mu_{s0}$	$\mu_{s1} > \mu_{s0}$
Interspersion and juxtaposition	$\mu_{s1} = \mu_{s0}$	$\mu_{s1} < \mu_{s0}$
Mean nearest-neighbour distance	$\mu_{s1} = \mu_{s0}$	$\mu_{s1} > \mu_{s0}$
Contagion index	$\mu_{s1} = \mu_{s0}$	$\mu_{s1} < \mu_{s0}$

μ_{s1} = mean of sample output at simulation step 20; μ_{s0} = value for initial data set

ventory data (Trudell 1996). The cover type layer is a raster representation of the dominant species within each polygon stand. The age of cover type layer represents the age of dominant species within each stand for the year the data were collected. The field capacity layer and the soil nutrient status layer were created from Ontario Land Inventory (OLI) data (OMNR 1977) and remain constant throughout model simulations. The field capacity layer indicates soil water holding capacity based on existing soil data. TerrAin was also used in creating the final field capacity layer to identify extreme wet and dry locations using a topographic wetness index (Wolock and McCabe 1995) calculated from a digital elevation model with 100-m horizontal and 1-m vertical resolution (NRC 2000). The soil nutrient status layer was derived from OLI soil texture data and represents soil nutrient availability. The geoclimatic zone layer was created by reclassifying Rowe's (1972) forest regions of Canada into a northern boreal mixedwood zone and a southern boreal mixedwood zone. These five layers were compiled for a study area (525,625 ha) in the boreal forest region of Ontario (Figure 1).

The VTM was applied with and without accounting for spatial dependence. The spatial dependence of the values being simulated was accounted for using spatial dependence of associated data layers to apply spatial bias. We refer to this method of applying spatial bias as the region approach because regions of influence are used to apply spatial bias to the assignment of probabilistic algorithms (region is a contiguous group of pixels with similar values). An advantage of the region approach is that spatial bias is applied variably in space depending on the strength and location of spatial dependence. Furthermore, regions can be identified based on spatial dependence of the layer being simulated or one or more of the associated data layers. Regions can also be defined using biological knowledge or ecological boundaries dependent on the model process and scale at which it operates. In this paper, we will demonstrate one way to create regions of spatial dependence, see Weaver (2002) for other methods.

Creating regions

For the case study, regions were created by quantifying the local spatial dependence of the field capacity and soil nutrient status layers. These layers were used because they are biologically important; soil conditions are associated closely with vegetation (Bonan and Shugart 1989;Prach et al. 1993;Harvey et al. 1995) and are key indicators of transition probabilities in the VTM.

Moran's *I* (Moran 1950), which is an indicator of spatial dependence, was used to quantify local spatial dependence of the two soil layers. For raster data, a Moran's *I* value of -1 represents perfect negative spatial dependence, 1 represents perfect positive spatial dependence, and $-1/(n-1)$, where n is the number of pixels, represents no spatial dependence. Moran's *I* was measured using a 5^2 pixel moving window in Blackland GRASS 2.0 (TAES 2001), with the resulting Moran's *I* value for each window assigned to the focal or centre pixel. Adjacency for measuring Moran's *I* was defined using a binary weight matrix (1 if adjacent, 0 otherwise) with diagonals excluded.

The two layers of Moran's *I* values, one for field capacity and one for soil nutrient status, were merged to form one layer representative of the underlying spatial dependence. A weighted average was used with the field capacity layer weighted higher than the soil nutrient status layer under the assumption that field capacity was more important in determining the transition probabilities. The resulting layer was then classified into non-positive (negative or no spatial dependence) and positive categories of spatial dependence. Only the positive spatial dependence regions were used to spatially bias transition probability assignment. It was assumed that the non-positive spatial dependence regions would exhibit negative or no spatial dependence intrinsically by not applying any spatial bias. The positive spatial dependence category was sub-divided into low-positive, intermediate-positive, and high-positive categories using an equal interval classification method. The classified layer was then grouped into regions based on contiguity (diago-

nals included) with identical values to provide each region with a unique identifier to define its membership and boundary. A spatial smoothing filter was applied to the region layer to filter out small regions (< 25 pixels) that were considered artifacts of the averaging and grouping processes. The final region layer was added as input into the VTM.

Probability assignment

The general process of assigning transitions was similar for the VTM whether it was applied with or without accounting for spatial dependence, however probabilities differed. The transition algorithm for the VTM without accounting for spatial dependence started by randomly selecting a pixel on the landscape. The transition probability vector corresponding to the cover type of that pixel was then selected using spatial information at that location (e.g., field capacity, age since last transition). Only pixels that were eligible for transition (i.e., holding times are variable) were selected. A random number generator was used to determine which transition cover type to assign that pixel and another pixel was selected at random on the landscape.

The transition algorithm that incorporates the region approach to account for spatial dependence started by randomly selecting a pixel and assigning a transition as outlined above. However, if the selected pixel was in a region of positive spatial dependence, then another pixel was selected from within only that region and the transition cover type assigned to the first pixel was used to bias all other probabilities in that region. By doing so, the cover type that was assigned in the first transition in the region had a higher probability of being assigned to other pixels in that region than it would have without spatial bias. All remaining pixels in that region were selected and assigned a transition based on the biased probabilities. If the randomly selected pixel was in a non-positive region, than no spatial bias was applied and the pixel was assigned a transition using the original probabilities. Figure 2 outlines the steps used in the spatial bias algorithm to account for spatial dependence in the VTM.

Probability weighting

Spatial bias was applied by weighting the transition probabilities for each pixel in a positive spatial dependence region using two equations. The first equation was used to bias the transition probability from the current cover type i to cover type j, with cover type j being the first assigned cover type in a region:

$$P_{ij} = P_{ij} + (1 - P_{ij})x, \qquad 0 = x = 1$$

where, P_{ij} is the biased probability for cover type j, P_{ij} is the original probability for cover type j, and x is the biasing factor. Constraints on the biasing factor ensured that the biased probability was never greater than one and the amount of weight applied was inversely proportional to the original probability. If the original probability was zero, no weight was applied to the transition probabilities. The degree of spatial dependence determined the value of the biasing factor: $x = 2/7$ for the low-positive, $x = 2/5$ for the intermediate-positive, and $x = 2/3$ for the high-positive regions. The biasing factors were arbitrarily chosen to test the effects of bias at different levels. The second equation was used to down-weight all other probabilities proportionally so that the remaining probabilities in the transition vector equaled $1 - P_{ij}$:

$$P_{ij_n} = \frac{P_{ij_n}}{(1 - P_{ij})} * (1 - P_{ij})$$

where, P_{ij_n} is the adjusted probability for cover type j_n, where $n = (1, 2, \ldots, m\text{-}1)$, P_{ij_n} is the original probability for cover type j_n, where $n = (1, 2, \ldots, m\text{-}1)$, P_{ij} is the original probability for cover type j, and P_{ij} is the biased probability for cover type j.

The VTM was applied for 25 simulation replicates for a period of 100 simulation steps, first without accounting for spatial dependence and then with accounting for spatial dependence. LPIs were used to quantify the spatial configuration for each simulated output at two simulation steps (50 and 100) for two extents (entire and local). For the local extent, the values for measured using 21 random windows were averaged prior to analysis. The LPI means for the output from two simulation scenarios (VTM with and without accounting for spatial dependence) were compared using separate pair-wise t-tests.

Results and discussion

Study 1: consequences of ignoring spatial dependence in LCTMs

The LPIs for the four distinct raster data sets converged as a result of using a LCTM without ac-

counting for spatial dependence, as only two of the 48 means were significantly different ($df = 98$, $p < 0.05$). Therefore, the first null hypothesis (no significant difference in spatial configurations among simulated output for four initially distinctly different data sets) could not be rejected. Convergence among data sets occurred at simulation step 20 and they remained converged through to simulation step 100. The exceptions were MPS and PD measured at the entire extent for output at simulation step 100 as they differed significantly among data sets. For those indices, convergence did occur, but in two separate subsets (block, small patch and large patch, random) as determined using Duncan's post-hoc test. Convergence of LPIs among distinct data sets has been reported before. Li and Reynolds (1997), using a Markov-cellular automata model, found that many of the LPIs they used to measure spatial pattern of different data sets converged after simulation.

Results of the *t*-tests revealed that the simulated output spatial configuration at simulation step 20 differed significantly from the initial data for the three non-random data sets with the exception of IJI and MNN (they could not be calculated because only one cover type occurred in each local window). All tests were significant, and therefore, the second null hypothesis (the simulated output spatial configuration of the non-random data sets are not significantly more fragmented than the initial data set) was rejected.

Plotting the means of each LPI vs. simulation steps revealed which data sets differed most from their initial spatial configuration and how the spatial configuration of the data sets changed through time (Figure 3). For the three non-random data sets the output at simulation step 20 was more fragmented than the initial data. For example, MPS, MNN, and CONTAG all decreased and PD decreased. The largest changes in spatial configuration from the initial conditions to the simulated outputs occurred in data sets with the most aggregated initial spatial configuration. Furthermore, the largest change in spatial configuration occurred by simulation step 20 with little change from simulation step 20 through to step 100 for all LPIs. All four data sets had a random spatial configuration with the identity of the initial data sets no longer evident by simulation step 20. The drastic change in the spatial configuration after 20 simulation steps was evident by comparing raster images of the initial block and large patch data sets with one replicate of their respective simulated output (Figure 4).

The results of fragmented output in this study agree with findings reported by others: Turner (1987) concluded that simulated landscapes became progressively fragmented because the transition model did not consider neighbourhood effects. Li and Reynolds (1997), in modelling rangeland degradation using a LCTM, concluded that all landscapes with initially different non-random patterns converged progressively to a random pattern.

Study 2: effects of accounting for spatial dependence in LCTMs

Comparisons between the simulated output from the VTM, with and without accounting for spatial dependence, revealed that outputs were significantly different (Table 2). Only two of 32 *t*-tests between output from the two simulation scenarios were not significantly different; PSCV at simulation step 50 measured at the local extent and DLFD at simulation step 50 measured at the entire extent. Therefore the third null hypothesis (no significant difference in spatial configuration between output simulated with and without accounting for spatial dependence) was rejected. Plots of the LPIs versus simulation steps provide strong evidence that the simulated output produced using the VTM while accounting for spatial dependence was less fragmented (i.e., more aggregated) (Figure 5). The indices that most clearly showed a less fragmented spatial configuration were PD, MPI, MNN, and CONTAG. In general, the indices suggested that the patches of forest cover types were larger, more contiguous, and closer together when spatial bias was applied in the VTM.

The difference in spatial configuration was evident from visually comparing raster images from the simulation scenarios; the output from simulations that accounted for spatial dependence contained larger aggregated patches (Figure 6; subsets of the study area are shown for clarity). In both subsets, patches of cover type B were larger and more contiguous, resulting in less overall fragmentation and more contiguous patches of cover type A when spatial bias was applied. The results are closer to what is typical of the boreal forest region of Ontario where a couple forest cover types dominate the landscape and patches are large and contiguous. For most LPIs, the difference between the initial data set and simulation output was smaller (i.e., less altered) for the output simulated with spatial bias. Exceptions were those indices that progressively became less fragmented for

Table 2. Summary of mean comparison results (output from simulations without accounting for spatial dependence vs. output from simulations accounting for spatial dependence) for landscape pattern indices

	Extent measured							
	Entire				Local			
	Simulation step				Simulation step			
	50		100		50		100	
Landscape pattern index	t	p	t	p	t	p	t	p
Mean patch size	− 176.1	< 0.0001	− 209.9	< 0.0001	− 22.0	< 0.0001	− 49.6	< 0.0001
Patch size coefficient of variation	3.5	0.0009	38.8	< 0.0001	1.8	0.0807	19.9	< 0.0001
Patch density	201.9	< 0.0001	262.3	< 0.0001	15.1	< 0.0001	32.9	< 0.0001
Double log fractal dimension	0.8	0.4050	14.5	< 0.0001	5.6	< 0.0001	21.8	< 0.0001
Mean proximity index	− 14.8	< 0.0001	− 35.8	< 0.0001	− 14.2	< 0.0001	− 27.7	< 0.0001
Interspersion and juxtaposition	69.4	< 0.0001	138.1	< 0.0001	15.6	< 0.0001	28.3	< 0.0001
Mean nearest-neighbour distance	− 73.9	< 0.0001	− 105.2	< 0.0001	− 12.1	< 0.0001	− 19.4	< 0.0001
Contagion index	− 82.9	< 0.0001	− 49.7	< 0.0001	− 11.9	< 0.0001	− 13.5	< 0.0001

n = 50, df = 48, numbers in italics are significant

both simulation scenarios (i.e., IJI and CONTAG). Those indices behaved differently due to their sensitivity to changing number of classes and changing proportions (Gustafson and Parker 1992;He et al. 2000; Baldwin et al. 2001). However, even in cases where both output types became less fragmented, the output simulated with accounting for spatial dependence remained less fragmented.

As displayed in Figure 5, indices such as MPS and MNN decreased while indices such as PSCV and PD increased from simulation step 0. Therefore, both with and without spatial bias simulate output became more fragmented than the initial spatial configuration. This is important because it shows that spatial bias as applied in this study using the region approach did not force artificial highly aggregated output. The region approach enables model users to limit, but not prohibit, fragmentation as a result of stochastic simulation.

Other researchers have also found that output became less fragmented when spatial bias was applied to a transition probability model. Turner (1987), Silvertown et al. (1992), and Wissel (1992) introduced spatial bias into a transition model using local neighbourhoods and found that the simulated output was less fragmented (i.e., highly aggregated). Similarly, Matsinos and Troumbis (2002) in modelling grassland dynamics using a CA model concluded that the initial random species distribution became clumped very quickly. However, many differences exist between the approach used in this study to account for spatial dependence and approaches used in other studies. Some studies (e.g., Turner 1987, Silvertown et al. 1992, Wissel 1992, Jenerette and Wu 2001, Matsinos and Troumbis 2002) have accounted for spatial dependence using various forms of CA models where neighbouring values are used to influence the transition probability of the focal pixel. These models assume homogeneity in spatial dependence because bias is applied uniformly across the modelled landscape. They do not incorporate spatial information (e.g., dependence) from other factors (e.g., soil data) that may influence the spatial dependence of the modelled values. Furthermore, neighbourhoods used to bias probabilities are often selected arbitrarily (Addicott et al. 1987), with no spatial or ecological basis.

We demonstrated a solution where the spatial bias comes from the spatial dependence of underlying data layers and is applied non-uniformly across the landscape; different amounts of bias were applied to various regions on the landscape based on the degree of spatial dependence. In addition, regions of influence were based on spatial properties of input data layers (i.e., field capacity and nutrients) that had an ecological association with forest cover type succession being modelled (e.g., Bonan and Shugart 1989, Prach et al. 1993, Harvey et al. 1995). This approach also provides flexibility in scale; the user can control the scale at which spatial dependence is measured and the scale of the final regions. Therefore, the user can adapt the region approach to a scale that is appropriate for the modelled process. As mentioned earlier, pixel resolu-

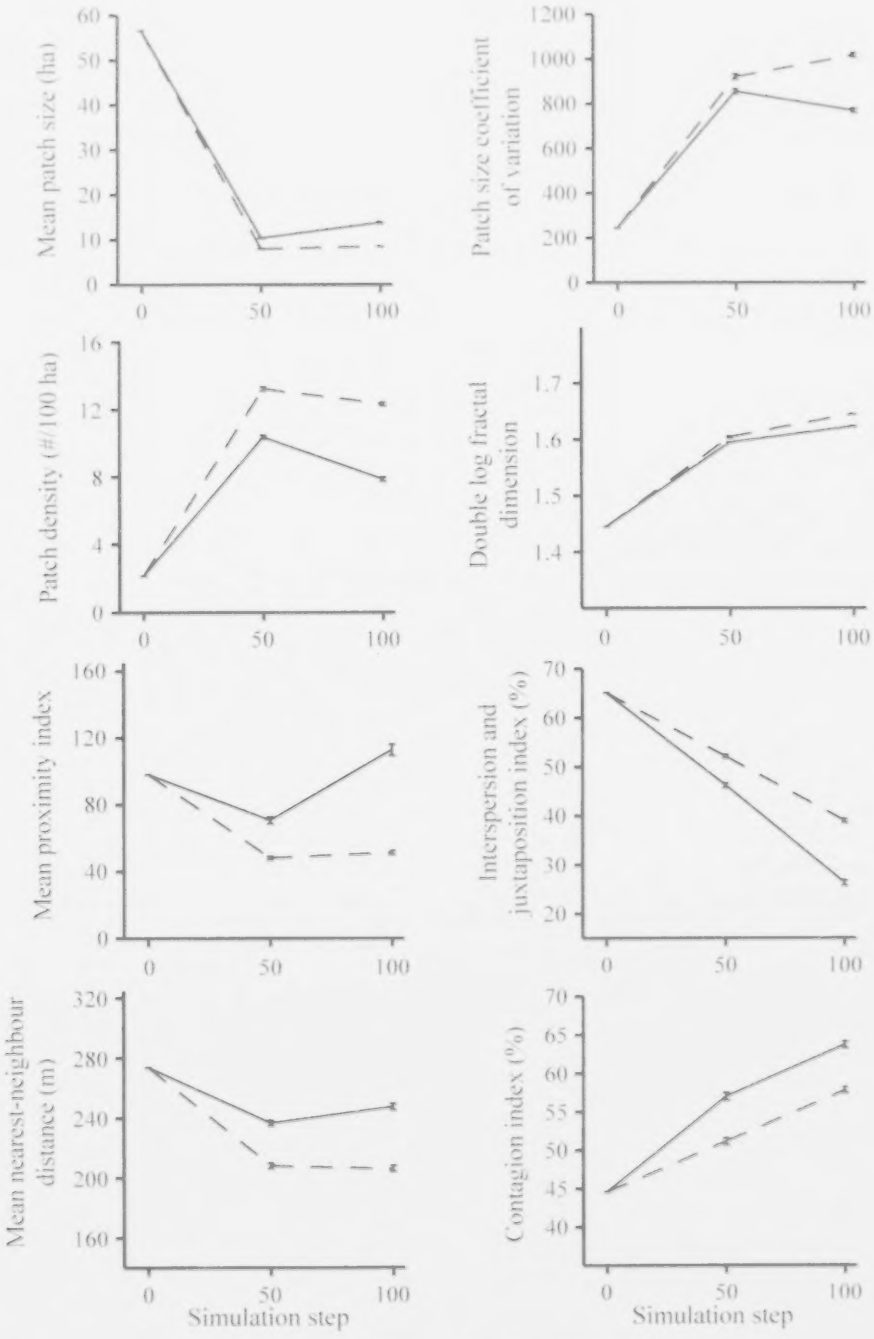

Figure 5. Means (± SE, n = 25) of landscape pattern indices (measured at the local extent) for simulated output from the VTM with and without accounting for spatial dependence.

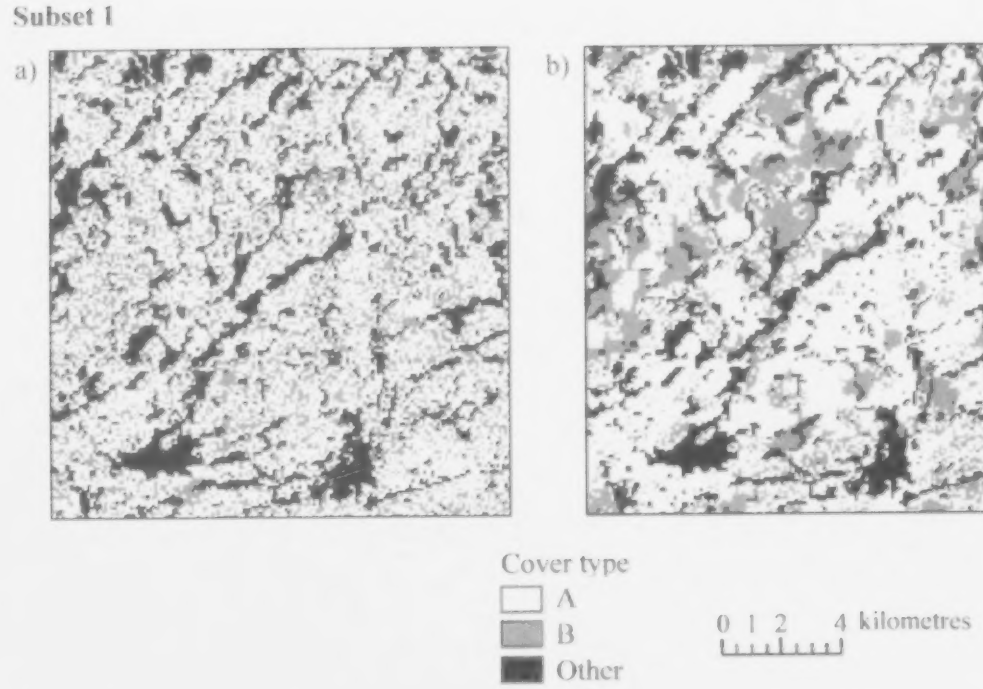

Figure 6. Two spatial subsets of simulated output (simulation step 100) from the Vegetation Transition Model a) without accounting for spatial dependence and b) with accounting for spatial dependence. Note: Simulation started from the same initial landscape.

tion of the input data will be a determining factor in the extent to which fragmentation may be a problem.

Conclusions

Even with a very simple model, the consequence of ignoring spatial dependence was clear. When probabilities were assigned aspatially, regardless of the in-

put spatial configuration, there was no difference in the output spatial configuration among four distinct data sets. Furthermore, the spatial configuration of the non-random data sets changed significantly, becoming more fragmented, and approached the spatial configuration of the random data set. When spatial dependence was ignored, the stochastic algorithm in the model artificially altered output spatial configuration indicating that spatial dependence should not be ignored. LCTMs that do not account for spatial dependence may provide output that is misleading, which in turn has implications for model users. Although the consequences of ignoring spatial dependence have been reported before, until now they have not been experimentally demonstrated. This study should increase awareness of the problem that some LCTMs ignore spatial dependence. It may prompt model developers to ensure that spatial properties are accounted for in the modelling process and serve as a caution for model users who assume that stochastic spatial models are spatially reliable.

We applied a region approach to account for spatial dependence in an existing LCTM. Based on the results, the region approach was effective as one possible solution to limiting fragmentation that occurred when spatial dependence was ignored. With the inclusion of spatial bias, spatial configuration of the simulated output was less fragmented. The purpose of the region approach was to limit artificial randomness, but not force a consistent or pre-determined spatial configuration. This was achieved since output from the VTM while accounting for spatial dependence remained less fragmented than when spatial dependence was ignored.

The specific results of this study were affected by the data, spatial scale, and methods used and should be interpreted with that in mind. Spatial dependence is an emergent property of the scale (resolution and extent) of the data and how it is quantified (i.e., Moran's I, 5^2 window). Furthermore, the spatial resolution, extent, and composition of the data (Turner et al. 1989; Baldwin et al. 2001) affect LPIs. Direct comparisons of LPIs with other studies should be conducted with caution.

The region approach provides a way to include spatial dependence in the modelling process based on a biological association with environmental variables, thus applying bias similar to what occurs in nature. Including spatial dependence adds realism to the model that may change the ecological outcome completely (Hunsaker et al. 2001). Furthermore, the region approach provides flexibility in accounting for spatial dependence and has potential for many types of ecological models. Regions can be defined in a way that makes ecological sense to the modelled process and the scale at which the regions are quantified and derived can be adjusted to match the scale of the data and processes.

Acknowledgements

We are thankful to many colleagues for assisting us: Ferenc Csillag for discussions on spatial dependence, Frank Schnekenburger for advice in programming, Dennis Yemshanov for modifying the Vegetation Transition Model for this study, David Baldwin for help in preparing spatial data, and Lisa Buse for editing an earlier draft of this manuscript. We are grateful for helpful comments from two anonymous reviewers.

References

Acevedo M.F., Urban D.L. and Shugart H.H. 1996. Models of forest dynamics based on roles of tree species. Ecological Modelling 87: 267–284.

Addicott J.F., Aho J.M., Antolin M.F., Padilla D.K., Richardson J.S. and Soluk D.A. 1987. Ecological neighbourhoods: scaling environmental patterns. Oikos 49: 340–346.

Anselin L. 1996. The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In: Fischer M., Scholten H.J. and Unwin D. (eds), Spatial Analytical Perspectives on GIS. GISDATA IV edition. Taylor and Francis Ltd., Bristol, Pennsylvania, USA, pp. 111–125.

Baker W.L. 1989. A review of models of landscape change. Landscape Ecology 2: 111–133.

Baker W.L. and Cai Y. 1992. The r.le programs for multiscale analysis of landscape structure using the GRASS geographical information system. Landscape Ecology 7: 291–302.

Baldwin D.J.B., Weaver K., Schnekenburger F. and Perera A.H. 2001. Sensitivity of landscape pattern indices to spatial extent, data resolution and classification detail in the managed forest of Ontario. For. Res. Rep. No. 150, Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario, Canada.

Balzter H., Braun P.W. and Kohler W. 1998. Cellular automata models for vegetation dynamics. Ecological Modelling 107: 113–125.

Barringer T.H. and Robinson V.B. 1981. Stochastic models of cover class dynamics. In: Proceedings of the 15th International Symposium on Remote Sensing of the Environment. Environmental Research Institute-Michigan, Ann Arbor, Michigan, USA, pp. 125–144.

Bell E.J. and Hinojosa R.C. 1977. Markov analysis of land use change: continuous time and stationary processes. Socio-Economic Planning Sciences 11: 13–17.

Bonan G.B. and Shugart H.H. 1989. Environmental factors and ecological processes in boreal forests. Annual Review of Ecology and Systematics 20: 1–28.

Brown D.G., Pijanowski B.C. and Duh J.D. 2000. Modeling the relationships between land use and land cover on private lands in the Upper Midwest, USA. Journal of Environmental Management 59: 247–263.

Cliff A.D. and Ord J.K. 1973. Spatial Autocorrelation. Pion Ltd., London, UK.

Deutschman D.H., Bradshaw G.A., Childress W.M., Daly K.L., Grunbaum D., Pascual M., Schumaker N.H. and Wu J. 1993. Mechanisms of patch formation. In: Levin S.A., Powell T.M. and Steele J.H. (eds), Patch Dynamics. Springer-Verlag, Berlin, Germany, pp. 184–209.

Ding Y. and Fotheringham A.S. 1992. The integration of spatial analysis and GIS. Computers, Environment, and Urban Systems 16: 3–19.

ESRI. 1998. ArcInfo for Windows, Version 7.2.1. Environmental Systems Research Institute, Redlands, California, USA.

Forman R.T.T. and Godron M. 1986. Landscape Ecology. John Wiley and Sons: New York, New York, USA.

Fox J.C., Ades P.K. and Bi H. 2001. Stochastic structure and individual-tree growth models. Forest Ecology and Management 154: 261–276.

Frelich L.E. and Reich P.B. 1999. Neighbourhood effects, disturbance severity, and community stability in forests. Ecosystems 2: 151–166.

Gardner R.H. 1999. RULE: Map generation and a spatial analysis program. In: Klopatek J.M. and Gardner R.H. (eds), Landscape Ecological Analysis: Issues and Applications. Springer-Verlag: New York, New York, USA, pp. 280–303.

Green D.G. 1989. Simulated effects of fire, dispersal and spatial pattern on competition within forest mosaics. Vegetatio 82: 139–153.

Gustafson E.J. and Parker G.R. 1992. Relationships between landcover proportion and indices of landscape spatial pattern. Landscape Ecology 7: 101–110.

Hargis C.D., Bissonette J.A. and David J.L. 1998. The behavior of landscape metrics commonly used in the study of habitat fragmentation. Landscape Ecology 13: 167–186.

Harvey B.D., Leduc A. and Bergeron Y. 1995. Early postharvest succession in relation to site type in the southern boreal forest of Quebec. Canadian Journal of Forest Research 25: 1658–1672.

He H.S., DeZonia B.E. and Mladenoff D. 2000. An aggregation index (AI) to quantify spatial patterns of landscapes. Landscape Ecology 15: 591–601.

He H.S. and Mladenoff D. 1999. Spatially explicit and stochastic simulation of forest-landscape fire disturbance and succession. Ecology 80: 81–99.

Hunsaker C.T., Goodchild M.F., Friedl M.A. and Case T.J. 2001. Spatial Uncertainty in Ecology: Implications for Remote Sensing and GIS Applications. Springer-Verlag: New York, New York, USA.

Jenerette G.D. and Wu J. 2001. Analysis and simulation of land-use change in the central Arizona-Phoenix region, USA. Landscape Ecology 16: 611–626.

Levin S.A. 1976. Population dynamic models in heterogeneous environments. Annual Review of Ecology and Systematics 7: 287–310.

Li H. and Reynolds J.F. 1994. A simulation experiment to quantify spatial heterogeneity in categorical maps. Ecology 75: 2446–2455.

Li H. and Reynolds J.F. 1997. Modeling effects of spatial pattern, drought, and grazing on rates of rangeland degradation: A combined Markov and cellular automaton approach. In: Quattrochi D.A. and Goodchild M.F. (eds), Scale in Remote Sensing and GIS. Lewis Publishers: New York, New York, USA, pp. 211–230.

Matsinos Y.G. and Troumbis A.Y. 2002. Modeling competition, dispersal and effects of disturbance in the dynamics of a grassland community using a cellular automaton model. Ecological Modelling 149: 71–83.

McGarigal K. and Marks B. 1995. FRAGSTATS: Spatial pattern analysis program for quantifying landscape structure. General Technical Report. PNW-GTR-351, United States Department of Agriculture Forest Service, Pacific Northwest Research Station, Portland, Oregon, USA.

Molofsky J., Bever J.D. and Antonovics J. 2001. Coexistence under positive frequency dependence. Proceedings of the Royal Society of London B 268: 273–277.

Moore A.D. 1990. The semi-Markov process: A useful tool in the analysis of vegetation dynamics for management. Journal of Environmental Management 30: 111–130.

Moran P.A.P. 1950. Notes on continuous stochastic phenomena. Biometrika 37: 17–23.

Muller M.R. and Middleton J. 1994. A Markov model of land use change dynamics in the Niagara Region, Ontario, Canada. Landscape Ecology 9: 151–157.

Natural Resources Canada (NRC) 2000. Canadian digital elevation data: Standards and specifications. Centre for Topographic Information Customer Support Group, Natural Resources Canada, Sherbrooke, Quebec, Canada.

O'Neill R.V., Hunsaker C.T., Timmins S.P., Jackson B.L., Jones K.B., Riitters K.H. and Wickham J.D. 1996. Scale problems in reporting landscape pattern at the regional scale. Landscape Ecology 11: 169–180.

Ontario Ministry of Natural Resources (OMNR) 1977. A ready reference. Ontario Land Inventory (A guide to the reading and understanding of the maps produced by the Ontario Land Inventory). Ontario Ministry of Natural Resources, Toronto, Ontario, Canada.

Perera A.H., Baldwin D.J.B. and Schnekenburger F. 1997. LEAP II: A landscape ecological analysis package for land use planners and managers. Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario, Canada. Forest Research Report No. 146.

Perera A.H., Yemshanov D., Schnekenburger F., Weaver K., Baldwin D.J.B. and Boychuck D. 2003. Boreal FOrest Landscape Dynamics Simulator (BFOLDS): A grid-based spatially stochastic model for predicting crown fire regime and forest cover transition. Forest Research Report No. 152, Ontario Ministry of Natural Resources, Ontario Forest Research Institute, Sault Ste. Marie, Ontario, Canada.

Prach K., Pysek P. and Smilauer P. 1993. On the rate of succession. Oikos 66: 343–346.

Riitters K.H., O'Neill R.V., Hunsaker C.T., Wickham J.D., Yankee D.H., Timmins S.P., Jones K.B. and Jackson B.L. 1995. A factor

analysis of landscape pattern and structure metrics. Landscape Ecology 10: 23–39.

Rossi R.E., Mulla D.J., Journel A.G. and Franz E.H. 1992. Geostatistical tools for modeling and interpreting ecological spatial dependence. Ecological Monographs 62: 277–314.

Rowe J.S. 1972. Forest regions of Canada. Canadian Forest Service, Department of the Environment, Ottawa, Ontario, Canada. Publication No. 1300.

Silvertown J., Holtier S., Johnson J. and Dale P. 1992. Cellular automation models of interspecific competition for space – the effect of pattern on process. Journal of Ecology 80: 527–534.

Sirakoulis G.Ch., Karafyllidis I. and Thanailakis A. 2000. A cellular automaton model for the effects of population movement and vaccination on epidemic propagation. Ecological Modelling 133: 209–233.

Sklar F.H. and Costanza R. 1991. The development of dynamic spatial models for landscape ecology: A review and prognosis. In: Turner M.G. and Gardner R.H. (eds), Quantitative Methods in Landscape Ecology. Springer-Verlag: New York, New York, USA, pp. 239–288.

Stocks C.E. and Wise S. 2000. The role of GIS in environmental modelling. Geographical and Environmental Modelling 4: 219–235.

Svenning J.-C. and Skov F. 2002. Mesoscale distribution of understorey plants in temperate forest (Kalo, Denmark): the importance of environment and dispersal. Plant Ecology 160: 169–185.

TAES 2001. Blackland GRASS 2.0. Texas Agricultural Experiment Station, Blackland Research Center, Texas A&M University, Temple, Texas, USA.

Trudell A. 1996. Forest Resource Inventory manual. Ontario Ministry of Natural Resources, Peterborough, Ontario, Canada.

Turner M.G. 1987. Spatial simulation of landscape changes in Georgia: A comparison of 3 transition models. Landscape Ecology 1: 29–36.

Turner M.G. 1989. Landscape ecology: The effects of pattern on process. Annual Review of Ecology and Systematics 20: 171–197.

Turner M.G., O'Neill R.V., Gardner R.H. and Milne B.T. 1989. Effects of changing spatial scale on the analysis of landscape pattern. Landscape Ecology 3: 153–162.

Weaver K. 2002. Spatial bias in landscape ecological simulations: A case study of accounting for spatial dependence in a raster-based stochastic model using a region approach. M.Sc. Thesis, Department of Geography, University of Toronto, Toronto, Ontario.

Wiens J.A. 1976. Population responses to patchy environments. Annual Review of Ecology and Systematics 7: 81–120.

Wissel C. 1992. Modelling the mosaic cycle of a Middle European beech forest. Ecological Modelling 63: 29–43.

Wolfram S. 1984. Cellular automata as models of complexity. Nature 311: 419–424.

Wolock D.M. and McCabe G.J. 1995. Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. Water Resources Research 31: 1315–1324.

Yemshanov D. and Perera A.H. 2002. A spatially explicit stochastic model to simulate boreal forest cover transitions: General structure and properties. Ecological Modelling 150: 189–209.